Comput. Optim. Appl. manuscript No.
(will be inserted by the editor)

Quantitative analysis for a class of two-stage stochastic linear variational inequality problems

Jie Jiang · Xiaojun Chen · Zhiping Chen

Abstract This paper considers a class of two-stage stochastic linear variational inequality problems whose first stage problems are stochastic linear box-constrained variational inequality problems and second stage problems are stochastic linear complementary problems having a unique solution. We first give conditions for the existence of solutions to both the original problem and its perturbed problems. Next we derive quantitative stability assertions of this two-stage stochastic problem under total variation metrics via the corresponding residual function. Moreover, we study the discrete approximation problem. The convergence and the exponential rate of convergence of optimal solution sets are obtained under moderate assumptions respectively. Finally, through solving a non-cooperative game in which each player's problem is a parameterized two-stage stochastic program, we numerically illustrate our theoretical results.

Keywords two-stage stochastic variational inequality · quantitative stability · discrete approximation · exponential convergence · non-cooperative game

Xiaojun Chen's work was partially supported by Hong Kong Research Grant Council grant PolyU153000/17P.
Zhiping Chen's work was partially supported by the National Natural Science Foundation of China Nos. 11991023, 11991020, 11735011.

Jie Jiang
College of Mathematics and Statistics, Chongqing University, Chongqing, China & Department of Applied Mathematics, The Hong Kong Polytechnic University, Hong Kong, China
E-mail: jie.jiang@connect.polyu.hk

Xiaojun Chen
Department of Applied Mathematics, The Hong Kong Polytechnic University, Hong Kong, China
E-mail: maxjchen@polyu.edu.hk

Zhiping Chen
School of Mathematics and Statistics, Xi'an Jiaotong University, Shaanxi, China
E-mail: zchen@xjtu.edu.cn

1 Introduction

We consider a class of two-stage stochastic linear variational inequality problems in the following form [2,3,5]:

$$\begin{cases} 0 \in Ax + \mathbb{E}_P[B(\xi)y(\xi)] + q_1 + \mathcal{N}_{[l,u]}(x), \\ 0 \le y(\xi) \perp M(\xi)y(\xi) + N(\xi)x + q_2(\xi) \ge 0, \text{ for } P\text{−a.e. } \xi \in \Xi, \end{cases} \tag{1}$$

where $l, u \in \mathbb{R}^n$ and $l < u$ in the sense of componentwise; $\xi : \Omega \to \Xi \subseteq \mathbb{R}^s$ with the probability space being $(\Omega, \mathcal{F}, \mathbb{P})$, $P \in \mathcal{P}(\Xi)$ and $\mathcal{P}(\Xi)$ denotes the set of probability distributions on the support set Ξ, $A \in \mathbb{R}^{n\times n}, q_1 \in \mathbb{R}^n$; $B(\cdot) : \mathbb{R}^s \to \mathbb{R}^{n\times m}$, $M(\cdot) : \mathbb{R}^s \to \mathbb{R}^{m\times m}$, $N(\cdot) : \mathbb{R}^s \to \mathbb{R}^{m\times n}$ and $q_2(\cdot) : \mathbb{R}^s \to \mathbb{R}^m$ are all matrix-valued or vector-valued mappings. The mathematical expectation $\mathbb{E}_P$ is taken in a componentwise fashion with respect to (w.r.t.) the corresponding probability distribution $P := \mathbb{P} \circ \xi^{-1}$. Problem (1) aims to find a pair $(x, y(\cdot)) \in [l, u] \times \mathcal{Y}$ satisfying (1), where $\mathcal{Y}$ is the collection of measurable functions from Ξ to $\mathbb{R}^m$ such that the expectation in the first stage problem of model (1) is well-defined. $\mathcal{N}_{[l,u]}(x)$ denotes the normal cone to the box $[l, u]$ at x. We say that problem (1) satisfies the relatively complete recourse if the second stage problem of (1) has a solution $y^*(x, \xi)$ for any $x \in [l, u]$ and a.e. every $\xi \in \Xi$.

The deterministic variational inequality problem has been extensively investigated, see monographs [6,7,10] and the references therein. Recently, to describe uncertainty in the complex decision process, stochastic variational inequality problems have been put forward and studied increasingly. Chen, Pong and Wets [1] introduced the two-stage stochastic variational inequality problem and an expected residual minimization procedure for solving it. Rockafellar and Wets [20] considered the multi-stage stochastic variational inequality problem when the support set is discrete, which lays a theoretical foundation for numerical solution by reformulating the multi-stage stochastic problem in an extensive form. Closely following this work, Rockafellar and Sun employed in [17] the well-known Progressive Hedging Method (PHM) to solve the multi-stage stochastic variational inequality problem. It is worth pointing out that PHM was introduced by Rockafellar and Wets in [18] to solve multi-stage stochastic programs. Recently, Chen, Sun and Xu [3] proposed a discretization scheme for the two-stage stochastic linear complementarity problem (SLCP) with continuous random variables, and studied the distributionally robust counterpart of the two-stage SLCP when the ambiguity set is constructed with the first order moment information. More recently, Chen, Shapiro and Sun [2] generalized the two-stage stochastic variational inequality problem to two-stage stochastic generalized equations. They studied the convergence of sample average approximation (SAA) without the relatively complete recourse assumption. As a special case, they also considered a mixed two-stage stochastic nonlinear variational inequality problem and examined the uniqueness of its solution and the exponential convergence of its discrete approximation.

It is easy to verify that the first stage problem of (1) can be equivalently rewritten as

$$x - \mathrm{mid}\{l, u, x - Ax - \mathbb{E}_P[B(\xi)y(\xi)] - q_1\} = 0,$$

where the "mid" function is defined componentwise as follows:

$$\mathrm{mid}\{l_i, u_i, z_i\} = \begin{cases} l_i, \ z_i < l_i, \\ z_i, \ l_i \le z_i \le u_i, \quad i = 1, \ldots, n. \\ u_i, \ z_i > u_i, \end{cases}$$

Assume that for any pair $(x, \xi) \in X \times \Xi$, the second stage SLCP of problem (1) has a unique solution $y^*(x, \xi)$. Then substituting it into the first stage problem, we obtain

$$0 \in Ax + \mathbb{E}_P[B(\xi)y^*(x, \xi)] + q_1 + \mathcal{N}_{[l,u]}(x),$$

where the right-hand side only depends on x. This inspires us to consider a residual function $f_P : \mathbb{R}^n \to \mathbb{R}_+$ as follows:

$$f_P(x) : = \|x - \mathrm{mid}\{l, u, x - Ax - \mathbb{E}_P[B(\xi)y^*(x, \xi)] - q_1\}\|^2. \tag{2}$$

If there is an $x \in \mathbb{R}^n$ such that $f_P(x) = 0$, then x must be a solution to problem (1). For the convenience of further discussion in the sequel, we equivalently consider the following box-constrained optimization problem

$$\min_{x \in [l,u]} \quad f_P(x). \tag{3}$$

In this paper, we analyze the quantitative stability of problem (1) by employing the minimization problem (3). It is noteworthy that recasting the stochastic variational inequality problem (1) as a stochastic (nonconvex) optimization problem, such as (3), provides a vehicle for conducting the analysis in this paper. It is not a necessarily avenue to compute an approximation solution, see for example [3, 9, 11, 20].

The main contributions of this paper can be summarized as follows. First, we examine different sufficient conditions for the existence of solutions to problem (1) and its perturbed problems. Next, under the assumption that the solution to the second stage problem is unique, we carry out quantitative stability analysis of problem (1) w.r.t. suitable probability metrics. Moreover, we consider the discrete approximation to problem (1), and derive both the convergence and exponential rate of convergence of the optimal solution sets of the discrete approximation problems to that of the original problem. Finally, to confirm these theoretical results as well as their applications, we consider a multi-player non-cooperative two-stage stochastic game problem and present numerical results by using PHM.

Throughout this paper, we adopt the following notation. $\mathbb{R}^n_+ := \{x \in \mathbb{R}^n : x \ge 0\}$ and $\mathbb{R}^n_{++} := \{x \in \mathbb{R}^n : x > 0\}$. $\mathbb{B}$ denotes the closed ball centred at zero with radius one in the corresponding space according to context. $\|\cdot\|$ stands for the Euclidean norm of a vector or the induced matrix norm. $\mathcal{P}_k(\Xi) :=$

$\{P \in \mathcal{P}(\Xi) : \mathbb{E}_P[\|\xi\|^k] < +\infty\}$. For any $a \in \mathbb{R}^n$ and $\mathcal{A}, \mathcal{B} \subseteq \mathbb{R}^n$, we define $d(a, \mathcal{B}) = \inf_{b \in \mathcal{B}} \|a - b\|$ and $d(\mathcal{A}, \mathcal{B}) = \sup_{a \in \mathcal{A}} \inf_{b \in \mathcal{B}} \|a - b\|$.

The paper is organized as follows. In section 2, we present the existence of solutions and quantitative stability results. In section 3, we consider the discrete approximation and conduct convergence analyses using problem (3). In section 4, we consider a multi-player non-cooperative two-stage stochastic game problem and its numerical process by PHM. In section 5, we make some concluding remarks.

2 Quantitative stability

Stability analysis of stochastic optimization problems is important for not only theoretical study but also numerical approximation. When we handle a stochastic optimization problem numerically, usually the first step is the discrete or empirical approximation to the included high dimensional integrals. Then some critical questions arise: what is the quantitative relationship between the original continuous problem and its discrete approximation; whether the optimal value and/or optimal solution set of the approximation problem converge to those of the original problem. All these questions can be answered through stability analysis. In view of this, we carry out the quantitative stability analysis of problem (1) in this section. For this purpose, we first introduce some prerequisites.

2.1 Prerequisites

Probability metrics are distance functions on the space of probability measures or probability distributions. In this paper, we need the so-called pseudo metric between two probability measures/distributions. We call them pseudo metrics because they usually do not satisfy the axioms of distance. In pseudo metrics, there is a large class of probability metrics called ζ-structure metrics.

Definition 1 (probability metric with ζ-structure, see [12]) Let $\mathcal{G}$ be a collection of real-valued measurable functions on support set Ξ. Then, for any two probability measures $P, Q \in \mathcal{P}(\Xi)$, we call

$$\mathbb{D}_{\mathcal{G}}(P, Q) = \sup_{g \in \mathcal{G}} |\mathbb{E}_P[g(\xi)] - \mathbb{E}_Q[g(\xi)]|$$

the ζ-structure probability metric between P and Q induced by $\mathcal{G}$.

$\mathbb{D}_{\mathcal{G}}(P, Q)$ is a pseudo metric because $\mathbb{D}_{\mathcal{G}}(P, Q) = 0$ does not imply $P = Q$ unless $\mathcal{G}$ is rich enough. Obviously, we have the symmetry and triangle inequality for $\mathbb{D}_{\mathcal{G}}$. It is known from Definition 1 that different ζ-structure metrics can be derived through choosing different $\mathcal{G}$s. For example, if we take

$$\mathcal{G}_{TV} := \{g : \Xi \to \mathbb{R} : g \text{ is measurable and } \sup_{\xi \in \Xi} |g(\xi)| \leq 1\},$$

the resulting ζ-structure metric

$$\mathbb{D}_{TV}(P,Q) := \sup_{g\in\mathcal{G}_{TV}} |\mathbb{E}_P[g(\xi)] - \mathbb{E}_Q[g(\xi)]|$$

is called the total variation metric. If

$$\mathcal{G}_{FM_p} := \left\{ g : \varXi \to \mathbb{R} : |g(\xi_1) - g(\xi_2)| \le \max\left\{1, \|\xi_1\|, \|\xi_2\|\right\}^{p-1} \|\xi_1 - \xi_2\| \right\},$$

the corresponding ζ-structure metric

$$\zeta_p(P,Q) := \sup_{g\in\mathcal{G}_{FM_p}} |\mathbb{E}_P[g(\xi)] - \mathbb{E}_Q[g(\xi)]|$$

is called the pth order Fortet-Mourier metric, which is often used in the stability analysis of stochastic programs. The selection of probability metric depends on the properties of the stochastic optimization problem. For example, to employ the total variation metric, some boundedness properties of the objective function are needed. This can be easily observed from its definition. The Fortet-Mourier metric requires some locally Lipschitz continuity conditions for the objective function, which is widely used in the quantitative stability analysis of two-stage stochastic linear programming problems. One can refer to [21] and [25] and references therein for more details. Here we employ these two ζ-structure metrics due to the boundedness and locally Lipschitz continuity of the corresponding objective functions. As for their equivalence, weak convergence and discrete approximations of above pseudo metrics, the reader is referred to [21].

In what follows, we give some useful properties about the solution to the second stage SLCP problem. A matrix $M \in \mathbb{R}^{m\times m}$ is a $\mathbb{P}$-matrix if all principal minors of M are positive.

Proposition 1 ([4]) *Let $M(\xi)$ be a $\mathbb{P}$-matrix for every $\xi \in \varXi$. The following assertions hold for problem* (1).

(i) For any given $x \in [l,u]$ and $\xi \in \varXi$, the second stage problem of (1) *has a unique solution $y^*(x,\xi)$, which can be implicitly written as*

$$y^*(x,\xi) = -W(x,\xi)(N(\xi)x + q_2(\xi)),$$

where $W(x,\xi) := [I - D(x,\xi)(I - M(\xi))]^{-1}D(x,\xi)$ and $D(x,\xi)$ is the m-dimensional diagonal matrix defined by

$$D_{jj}(x,\xi) = \begin{cases} 1, & \text{if } (M(\xi)y^*(x,\xi) + N(\xi)x + q_2(\xi))_j \le y_j^*(x,\xi), \\ 0, & \text{otherwise} \end{cases}$$

for $j = 1,\cdots,m$;

(ii) $y^*(\cdot,\xi)$ is Lipschitz continuous, i.e.,

$$\|y^*(x_1,\xi) - y^*(x_2,\xi)\| \le \max_{J\in\mathcal{J}} \left\|M_{J\times J}^{-1}(\xi)\right\| \|N(\xi)\| \|x_1 - x_2\|,$$

where $M_{J\times J}(\xi)$ is the sub-matrix of $M(\xi)$, whose entries are indexed by $J\times J$, and $\mathcal{J}$ denotes the power set of $\{1,2,\cdots,m\}$.

For further discussion, we need the following assumption (see, for example, [2,3]).

Assumption 1 *Let $M(\xi)$ be a $\mathbb{P}$-matrix for every $\xi \in \Xi$. Moreover, there exists a continuous function $\kappa_M : \Xi \to \mathbb{R}_{++}$, such that*

$$\max_{J \in \mathcal{J}} \left\| M_{J\times J}^{-1}(\xi) \right\| \leq \frac{1}{\kappa_M(\xi)}$$

for any $\xi \in \Xi$.

A sufficient condition for Assumption 1 is $y^T M(\xi) y \geq \kappa_M(\xi) \|y\|^2$ for any $y \in \mathbb{R}^m$ and $\xi \in \Xi$, from which we can deduce from [3, Lemma 2.1] that $M(\xi)$ is a $\mathbb{P}$-matrix and in addition $\left\| M_{J\times J}^{-1}(\xi) \right\| \leq \frac{1}{\kappa_M(\xi)}$ for any $J \in \mathcal{J}$. A stronger assumption is adopted in [3, Assumption 2.1] (see Assumption 2 below).

2.2 Existence of solutions

The existence of solutions to stochastic variational inequality problems have been studied in [8,13–16]. Specially, Ravat and Shanbhag considered in [13] the stochastic Nash game where the expectation of each player's cost function is minimized. Conditions to admit an equilibrium for both smooth and nonsmooth (but continuous) objective functions were investigated. More recently, the same authors discussed in [14] some verifiable sufficiency conditions for the existence of solutions to stochastic (quasi-)variational inequality problems which extended the results in [13] from single-valued stochastic variational inequality problems to multi-valued stochastic quasi-variational inequality problems.

The existing works mainly concentrate on the deterministic case or the single-stage case. Here, we adopt these pioneering works or concepts to give some assertions about the existence of solutions to the two-stage stochastic ones. In the two-stage case, Chen, Sun and Xu employed the strong monotonicity in terms of a redefined inner product on the product space of the first stage and second stage variables in [3], under which the existence and uniqueness assertion of solutions to the two-stage SLCP were derived. Under Assumption 1, we know that there always exists a unique solution $y^*(x, \xi)$ to the second stage SLCP for any given pair $(x, \xi) \in [l, u] \times \Xi$. Namely, problem (1) satisfies the relatively complete recourse condition. However, this does not necessarily ensure the existence of a solution to problem (1). Therefore, in the sequel, we will introduce several conditions such that problem (1) has at least one solution under probability distribution P, and so does its perturbed problem under Q, i.e.,

$$\begin{cases} 0 \in Ax + \mathbb{E}_Q[B(\xi)y(\xi)] + q_1 + \mathcal{N}_{[l,u]}(x), \\ 0 \leq y(\xi) \perp M(\xi)y(\xi) + N(\xi)x + q_2(\xi) \geq 0, \text{ for } Q\text{--}a.e.\ \xi \in \Xi. \end{cases} \tag{4}$$

To introduce the first sufficient condition, we make the following assumption which was first used in [3] to study the two-stage SLCP.

Assumption 2 *There exists a continuous function $\kappa(\cdot) : \Xi \to \mathbb{R}_{++}$, such that*

$$(x^T, y^T) \begin{pmatrix} A & B(\xi) \\ N(\xi) & M(\xi) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \geq \kappa(\xi)(\|x\|^2 + \|y\|^2) \tag{5}$$

P-a.e. $\xi \in \Xi$, for any $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$, where $\mathbb{E}_P[\kappa(\xi)] < +\infty$.

It is easy to see that Assumption 2 implies Assumption 1 by letting $x = 0$. Then, under Assumption 2, Chen, Sun and Xu [3] gave the following conclusion.

Proposition 2 *Suppose that Assumption 2 holds. Then problem* (1) *has a unique solution.*

Assumption 2 is sufficient for problem (1) to have a unique solution. In this paper, we give a weaker condition for the existence of solutions to problem (1) without the uniqueness. For this purpose, we introduce the following notation and the concept of pseudomonotonicity.

Define the mapping $\Phi_P : \mathbb{R}^n \to \mathbb{R}^n$ as

$$\Phi_P(x) = Ax + \mathbb{E}_P[B(\xi)y^*(x,\xi)] + q_1.$$

Recall that Φ_P is pseudomonotone [7, Definition 2.3.1] if

$$\langle x_1 - x_2, \Phi_P(x_2)\rangle \geq 0 \Rightarrow \langle x_1 - x_2, \Phi_P(x_1)\rangle \geq 0.$$

Immediately, based on [7], we have the following proposition.

Proposition 3 *Suppose that Assumption 1 holds and the following integral*

$$\int_{\Xi} \|B(\xi)\| \max_{J\in\mathcal{J}} \left\|M^{-1}_{J\times J}(\xi)\right\| \|N(\xi)\| \, P(d\xi)$$

is finite. Then the solution set of problem (1) *is nonempty and its projection on the first stage variable is compact. If, in addition, Φ_P is pseudomonotone on $[l,u]$, this projection is convex too.*

Proof We first verify that $\Phi_P(x)$ is continuous w.r.t. x. Note that

$$\|\Phi_P(x_1) - \Phi_P(x_2)\| \leq \|A\| \, \|x_2 - x_1\| + \|\mathbb{E}_P[B(\xi)(y^*(x_2,\xi) - y^*(x_1,\xi))]\| \, . \tag{6}$$

For the second term of the right-hand side of (6), we have estimation

$$\begin{aligned} &\|\mathbb{E}_P[B(\xi)(y^*(x_2,\xi) - y^*(x_1,\xi))]\| \\ &\leq \mathbb{E}_P[\|B(\xi)(y^*(x_2,\xi) - y^*(x_1,\xi))\|] \\ &\leq \mathbb{E}_P[\|B(\xi)\| \max_{J\in\mathcal{J}} \left\|M^{-1}_{J\times J}(\xi)\right\| \|N(\xi)\|] \, \|x_1 - x_2\| \, . \end{aligned}$$

Due to the finiteness of $\mathbb{E}_P[\|B(\xi)\| \max_{J\in\mathcal{J}} \left\|M^{-1}_{J\times J}(\xi)\right\| \|N(\xi)\|]$, we know that Φ_P is Lipschitz continuous, which is obviously continuous. Then, we derive from [7, Proposition 2.2.3] that

$$-\Phi_P(x) \in \mathcal{N}_{[l,u]}(x) \tag{7}$$

has a solution. Since Assumption 1 holds, there always exists a solution for the second stage problem for any $x \in [l, u]$. To summarize, problem (1) has a solution.

Corollary 2.2.5 in [7] tells us that: if $X \subseteq \mathbb{R}^n$ is compact and convex, and $F : X \to \mathbb{R}^n$ is continuous, the solution set of $-F(x) \in \mathcal{N}_X(x)$ is nonempty and compact. If, in addition, F is pseudomonotone, it is known from [7, Theorem 2.3.5] that the solution set is convex.

Due to the boundedness and convexity of interval $[l, u]$, we know from [7, Corollary 2.2.5] that the solution set of (7) is nonempty and compact. Moreover, if Φ_P is pseudomonotone, based on [7, Theorem 2.3.5], the solution set of (7) is convex. □

To establish the existence of solutions to the perturbed problem, we need the following assumption.

Assumption 3 *There exist constants $\alpha \geq 0$ and $C > 0$, such that the random matrices and vector in problem* (1) *are bounded as*

$$\|\Lambda(\xi)\| \leq C \max\{1, \|\xi\|\}^{\alpha}, \text{ for a.e. } \xi \in \Xi,$$

where $\Lambda(\xi) = B(\xi), M(\xi), N(\xi)$ or $q_2(\xi)$.

Remark 1 We have the following observations for the assumptions in Proposition 3.

(i) If Assumptions 1 and 3 hold, and $\kappa_M(\xi) \geq \kappa$ for some positive constant κ, we have

$$\|B(\xi)\| \max_{J \in \mathcal{J}} \left\| M_{J \times J}^{-1}(\xi) \right\| \|N(\xi)\| \leq \frac{C^2 \max\{1, \|\xi\|\}^{2\alpha}}{\kappa}.$$

Then a sufficient condition for the integrability of the left-hand side is simply $P \in \mathcal{P}_{2\alpha}(\Xi)$, which can be verified easily.

(ii) As for the pseudomonotonicity of Φ_P over $[l, u]$, we can verify the monotonicity of Φ_P instead of the pseudomonotonicity if possible. This might be easier to implement, which is only necessary to examine the monotonicity of $Ax + B(\xi)y^*(x, \xi) + q_1$ for almost everywhere $\xi \in \Xi$. It is known from Proposition 1 that, for any $x_1, x_2 \in [l, u]$, we have

$$\begin{aligned}
&\langle x_1 - x_2, Ax_1 + B(\xi)y^*(x_1, \xi) + q_1 - (Ax_2 + B(\xi)y^*(x_2, \xi) + q_1)\rangle \\
&\quad \geq \langle x_1 - x_2, A(x_1 - x_2)\rangle - \max_{J \in \mathcal{J}} \left\| M_{J \times J}^{-1}(\xi) \right\| \|B(\xi)\| \|N(\xi)\| \|x_1 - x_2\|^2 \\
&\quad \geq \left(\lambda_{\min}(A) - \max_{J \in \mathcal{J}} \left\| M_{J \times J}^{-1}(\xi) \right\| \|B(\xi)\| \|N(\xi)\| \right) \|x_1 - x_2\|^2,
\end{aligned}$$

where $\lambda_{\min}(A)$ is the minimal eigenvalue of A. Then a sufficient condition for the monotonicity is that

$$\lambda_{\min}(A) - \max_{J \in \mathcal{J}} \left\| M_{J \times J}^{-1}(\xi) \right\| \|B(\xi)\| \|N(\xi)\| \geq 0$$

holds for a.e. $\xi \in \Xi$. This can be further simplified under some specific settings. For example, if we have $y^*(x,\xi) = -M(\xi)^{-1}(N(\xi)x + q_2(\xi)) \geq 0$ for almost everywhere $\xi \in \Xi$ and each $x \in [l,u]$, that is, $A - B(\xi)W(x,\xi)N(\xi) = A - B(\xi)M(\xi)^{-1}N(\xi)$, the monotonicity condition holds when $A - B(\xi)M(\xi)^{-1}N(\xi)$ is positive semidefinite for a.e. $\xi \in \Xi$, which can be easily verified.

In what follows, we consider the existence of solutions to the perturbed problem (4) under certain conditions. To ease the statement, we define the multifunction $\Theta_P : [l,u] \rightrightarrows \mathbb{R}^n$ as

$$\Theta_P(x) = Ax + \mathbb{E}_P[B(\xi)y^*(x,\xi)] + q_1 + \mathcal{N}_{[l,u]}(x)$$

and its inverse is

$$\Theta_P^{-1}(y) := \{x \in [l,u] : y \in \Theta_P(x)\}.$$

Proposition 4 *Under assumptions of Proposition 3, $z \in \Theta_P(x)$ is solvable for any $z \in \mathbb{R}^n$.*

Proof Note that, for any $x \in \mathbb{R}^n$, $z \in \Theta_P(x)$ is equivalent to $-\Phi_P(x) + z \in \mathcal{N}_{[l,u]}(x)$. We know from the proof of Proposition 3 that $-\Phi_P(x)$ is continuous w.r.t x, so does $-\Phi_P(x)+z$. Then, by the same argument as that in Proposition 3, we know that $z \in \Theta_P(x)$ is solvable for any $x \in \mathbb{R}^n$, which completes the proof. □

With the aid of Assumption 3, we have the following lemma.

Lemma 1 *Suppose that Assumptions 1 and 3 hold, $\kappa_M(\xi) \geq \kappa > 0$ and $P, Q \in \mathcal{P}_{2\alpha+1}(\Xi)$. Then there exists a positive number L such that*

$$\|\mathbb{E}_P[B(\xi)y^*(x,\xi)] - \mathbb{E}_Q[B(\xi)y^*(x,\xi)]\| \leq L\mathbb{D}_{TV}(P,Q)^{\frac{1}{2\alpha+1}}, \tag{8}$$

when $\mathbb{D}_{TV}(P,Q) + \zeta_{2\alpha+1}(P,Q) \leq 1$.

Proof From [4, Theorem 2.1], the matrix $W(x,\xi)$ defined in Proposition 1 is well-defined and

$$\|W(x,\xi)\| \leq \max_{J \in \mathcal{I}} \left\| M_{J\times J}^{-1}(\xi) \right\|.$$

This and Assumption 1 imply that $\|W(x,\xi)\| \leq \frac{1}{\kappa_M(\xi)}$. Under Assumptions 1 and 3, we have from (i) of Proposition 1 that

$$\|y^*(x,\xi)\| \leq \frac{1}{\kappa_M(\xi)} \|N(\xi)x + q_2(\xi)\| \leq \frac{(R+1)C}{\kappa} \max\{1, \|\xi\|\}^{\alpha}, \tag{9}$$

where $R := \max_{x\in[l,u]} \|x\|$. Thus,

$$\|B(\xi)y^*(x,\xi)\| \leq \frac{(R+1)C^2}{\kappa} \max\{1, \|\xi\|\}^{2\alpha}. \tag{10}$$

Meanwhile, we have

$$\begin{aligned}\|\mathbb{E}_P[B(\xi)y^*(x,\xi)]-\mathbb{E}_Q[B(\xi)y^*(x,\xi)]\| \le &\int_{\{\xi\in\Xi:\|\xi\|>\Gamma\}}\|B(\xi)y^*(x,\xi)\|\,(P+Q)(d\xi)\\ &+\left\|\int_{\{\xi\in\Xi:\|\xi\|\le\Gamma\}}B(\xi)y^*(x,\xi)(P-Q)(d\xi)\right\|.\end{aligned}$$

Here, we select $\Gamma\ge 1$. For the second term at the right-hand side, we have

$$\begin{aligned}&\left\|\int_{\{\xi\in\Xi:\|\xi\|\le\Gamma\}}B(\xi)y^*(x,\xi)(P-Q)(d\xi)\right\|\\ &\quad=\frac{(R+1)C^2\Gamma^{2\alpha}}{\kappa}\left\|\int_{\{\xi\in\Xi:\|\xi\|\le\Gamma\}}\frac{B(\xi)y^*(x,\xi)}{(R+1)C^2\Gamma^{2\alpha}/\kappa}(P-Q)(d\xi)\right\|.\end{aligned}$$

It is known from (10) that

$$\|B(\xi)y^*(x,\xi)\|\le\frac{(R+1)C^2\Gamma^{2\alpha}}{\kappa}$$

for any ξ with $\|\xi\|\le\Gamma$. This implies

$$\frac{(B(\xi)y^*(x,\xi))_i}{(R+1)C^2\Gamma^{2\alpha}/\kappa}\le 1$$

due to $|(B(\xi)y^*(x,\xi))_i|\le\|B(\xi)y^*(x,\xi)\|$, for $i=1,2,\cdots,n$. Define $g_i(x,\xi)$ by

$$g_i(x,\xi)=\begin{cases}\frac{(B(\xi)y^*(x,\xi))_i}{(R+1)C^2\Gamma^{2\alpha}/\kappa}, & \|\xi\|\le\Gamma;\\ 0, & \text{otherwise.}\end{cases}$$

Obviously, we have $g_i(x,\xi)\in\mathcal{G}_{TV}$, which indicates that

$$\left|\int_{\Xi}g_i(x,\xi)(P-Q)(d\xi)\right|\le\mathbb{D}_{TV}(P,Q)$$

for $i=1,2,\cdots,n$. Denote by $g=(g_1,\cdots,g_n)^T$. Then, by the definition of total variation metric, we have

$$\begin{aligned}\left\|\int_{\{\xi\in\Xi:\|\xi\|\le\Gamma\}}\frac{B(\xi)y^*(x,\xi)}{(R+1)C^2\Gamma^{2\alpha}/\kappa}(P-Q)(d\xi)\right\| &=\left\|\int_{\Xi}g(x,\xi)(P-Q)(d\xi)\right\|\\ &=\left(\sum_{i=1}^{n}\left|\int_{\Xi}g_i(x,\xi)(P-Q)(d\xi)\right|^2\right)^{\frac{1}{2}}\\ &\le\sqrt{n}\mathbb{D}_{TV}(P,Q).\end{aligned}$$

Finally, we obtain

$$\left\|\int_{\{\xi\in\Xi:\|\xi\|\le\Gamma\}}B(\xi)y^*(x,\xi)(P-Q)(d\xi)\right\|\le\sqrt{n}\frac{(R+1)C^2}{\kappa}\Gamma^{2\alpha}\mathbb{D}_{TV}(P,Q).$$

Note that $\|\xi\|^p/p \in \mathcal{G}_{FM_p}$ for any $p \geq 1$, which means

$$\int_{\Xi} \|\xi\|^p Q(d\xi) - \int_{\Xi} \|\xi\|^p P(d\xi) \leq p\zeta_p(P,Q).$$

Thus,

$$\begin{aligned}
&\int_{\{\xi\in\Xi:\|\xi\|>\Gamma\}} \|B(\xi)y^*(x,\xi)\| (P+Q)(d\xi) \\
&\overset{(10)}{\leq} \int_{\{\xi\in\Xi:\|\xi\|>\Gamma\}} \frac{(R+1)C^2}{\kappa} \max\{1,\|\xi\|\}^{2\alpha}(P+Q)(d\xi) \\
&\leq \frac{(R+1)C^2}{\kappa\Gamma} \int_{\{\xi\in\Xi:\|\xi\|>\Gamma\}} \|\xi\|^{2\alpha+1} (P+Q)(d\xi) \\
&\leq \frac{(R+1)C^2}{\kappa\Gamma} \left(2\mathbb{E}_P[\|\xi\|^{2\alpha+1}] + (2\alpha+1)\zeta_{2\alpha+1}(P,Q)\right).
\end{aligned}$$

To summarize the above estimation, we obtain that

$$\begin{aligned}
&\|\mathbb{E}_P[B(\xi)y^*(x,\xi)] - \mathbb{E}_Q[B(\xi)y^*(x,\xi)]\| \leq \\
&\qquad \sqrt{n}\frac{(R+1)C^2}{\kappa}\Gamma^{2\alpha}\mathbb{D}_{TV}(P,Q) + \frac{(R+1)C^2}{\kappa\Gamma}(2\mathbb{E}_P[\|\xi\|^{2\alpha+1}] + 2\alpha+1),
\end{aligned}$$

which comes from the assumption that $\mathbb{D}_{TV}(P,Q)+\zeta_{2\alpha+1}(P,Q) \leq 1$. Specially, we define

$$\Gamma = \mathbb{D}_{TV}(P,Q)^{-1/(2\alpha+1)} \geq 1.$$

Finally, we derive that

$$\|\mathbb{E}_P[B(\xi)y^*(x,\xi)] - \mathbb{E}_Q[B(\xi)y^*(x,\xi)]\| \leq L\mathbb{D}_{TV}(P,Q)^{\frac{1}{2\alpha+1}},$$

where $L = \left((R+1)C^2(\sqrt{n} + 2\mathbb{E}_P[\|\xi\|^{2\alpha+1}] + 2\alpha+1)\right)/\kappa$. □

Proposition 5 *Suppose that Assumptions 1 and 3 hold, $\kappa_M(\xi) \geq \kappa > 0$ a.s. and $P, Q \in \mathcal{P}_{2\alpha}(\Xi)$. Then the perturbed problem (4) is solvable.*

Proof We have from Assumptions 1 and 3, and $\kappa_M(\xi) \geq \kappa > 0$ almost surely (a.s.) that

$$\begin{aligned}
\|B(\xi)\| \max_{J\in\mathcal{J}} \left\|M^{-1}_{J\times J}(\xi)\right\| \|N(\xi)\| &\leq \frac{C^2\max\{1,\|\xi\|^{2\alpha}\}}{\kappa_M(\xi)} \\
&\overset{\text{a.s.}}{\leq} \frac{C^2\max\{1,\|\xi\|^{2\alpha}\}}{\kappa}.
\end{aligned}$$

Recall $\mathcal{P}_{2\alpha}(\Xi) = \{P \in \mathcal{P}(\Xi) : \mathbb{E}_P[\|\xi\|^{2\alpha}] < +\infty\}$. This and $P \in \mathcal{P}_{2\alpha}(\Xi)$ imply that Proposition 4 holds.

Moreover, it is known from (10) that

$$\|B(\xi)y^*(x,\xi)\| \overset{\text{a.s.}}{\leq} \frac{(R+1)C^2}{\kappa}\max\{1,\|\xi\|\}^{2\alpha}.$$

Since $P, Q \in \mathcal{P}_{2\alpha}(\Xi)$, we obtain that both $\mathbb{E}_P[B(\xi)y^*(x,\xi)]$ and $\mathbb{E}_Q[B(\xi)y^*(x,\xi)]$ are well-defined and have finite value. Let

$$z = \mathbb{E}_P[B(\xi)y^*(x,\xi)] - \mathbb{E}_Q[B(\xi)y^*(x,\xi)] \in \mathbb{R}^n.$$

According to Proposition 4,

$$\mathbb{E}_P[B(\xi)y^*(x,\xi)] - \mathbb{E}_Q[B(\xi)y^*(x,\xi)] \in \Theta_P(x)$$

is solvable, that is,

$$0 \in Ax + \mathbb{E}_Q[B(\xi)y^*(x,\xi)] + q_1 + \mathcal{N}_{[l,u]}(x)$$

or the perturbed problem (4) is solvable. □

King and Rockafellar in [8] put forward the concept of subinvertibility to investigate the existence of the solution to perturbed generalized equations, which can be applied to the situation without the differentiability assumption. In the following, we employ the concept of subinvertibility to establish the existence assertion. The subinvertibility of a multifunction is defined on its graph. For more details about the graph of a multifunction, one can refer to [19]. Specifically, we have the following definition of subinvertibility.

Definition 2 (subinvertibility, [8]) $\Theta_P(x)$ is said to be subinvertible at $(x^*, 0)$, if $0 \in \Theta_P(x^*)$ and there exist a compact neighborhood U of x^*, a positive scalar ϵ and a nonempty convex-valued multifunction $G : \epsilon\mathbb{B} \to U$, such that the graph of G, denoted by gph$G$, is closed, $x^* \in G(0)$ and $G(y)$ is contained in $\Theta_P^{-1}(y)$ for each $y \in \epsilon\mathbb{B}$.

As for more discussion of subinvertibility, one can refer to [8] for details. Then, based on the concept of subinvertibility and [8, Proposition 3.1], we have the following proposition.

Proposition 6 *Suppose that all assumptions in Lemma 1 hold and $\Theta_P(x)$ is subinvertible at $(x^*, 0)$. Then there exist a compact and convex neighborhood U of x^* and a scalar $\epsilon \in (0, 1]$, such that*

$$0 \in Ax + \mathbb{E}_Q[B(\xi)y^*(x,\xi)] + q_1 + \mathcal{N}_{[l,u]}(x)$$

has at least one solution in U for every $Q \in \mathcal{P}_{2\alpha}(\Xi)$ satisfying $\mathbb{D}_{TV}(P,Q) + \zeta_{2\alpha+1}(P,Q) \le \epsilon$.

Proof According to Lemma 1, if $\mathbb{D}_{TV}(P,Q) + \zeta_{2\alpha+1}(P,Q) \le 1$, there is $L > 0$ such that

$$\|\mathbb{E}_Q[B(\xi)y^*(x,\xi)] - \mathbb{E}_P[B(\xi)y^*(x,\xi)\| \le L\mathbb{D}_{TV}(P,Q)^{\frac{1}{2\alpha+1}}.$$

From [8, Proposition 3.1], we know that there exists an $\epsilon_0 > 0$ such that the perturbed problem

$$\begin{aligned}
0 &\in Ax + \mathbb{E}_Q[B(\xi)y^*(x,\xi)] + q_1 + \mathcal{N}_{[l,u]}(x) \\
&= Ax + \mathbb{E}_P[B(\xi)y^*(x,\xi)] + q_1 + \mathcal{N}_{[l,u]}(x) \\
&\quad + Ax + \mathbb{E}_Q[B(\xi)y^*(x,\xi)] + q_1 - (Ax + \mathbb{E}_P[B(\xi)y^*(x,\xi)] + q_1)
\end{aligned}$$

has at least one solution in a compact and convex neighborhood U of x^* for any Q satisfying

$$L\mathbb{D}_{TV}(P,Q)^{\frac{1}{2\alpha+1}} \leq \epsilon_0.$$

Let $\epsilon = \min\{\left(\frac{\epsilon_0}{L}\right)^{2\alpha+1}, 1\}$. We complete the proof. □

The subinvertibility of $\Theta_P(x)$ can be verified under some typical cases, see [8]. The following remark tells us that our conditions are not limiting compared with those in [3,17].

Remark 2 As we mentioned before, in [3], the authors required Assumption 2, which is stronger than Assumption 1. On the other hand, in [17], the authors directly assumed that problem (1) is solvable, and the coefficient matrix

$$\begin{pmatrix} A & B(\xi) \\ N(\xi) & M(\xi) \end{pmatrix}$$

is positive semidefinite for any $\xi \in \Xi$, where their support set Ξ is assumed to be finite. In this case, the positive semidefinite assumption is equivalent to monotonicity. Our conditions are weaker than those in [17]. To clarify this, we consider the following coefficient matrix:

$$\begin{pmatrix} A & 0 \\ N \cdot \xi & M \cdot \xi^2 \end{pmatrix},$$

where $\xi \in \Xi := [\frac{1}{2}, 1]$, $A \in \mathbb{R}^{n\times n}$ is negative definite, $N \in \mathbb{R}^{m\times n}$ and $M \in \mathbb{R}^{m\times m}$ is positive definite. Obviously, due to the negative definiteness of A, this kind of coefficient matrix is not positive semidefinite. When the coefficient matrix takes the above form, the first stage problem is always solvable if $x = -A^{-1}q_1 \in [l, u]$. Moreover, the positive semidefiniteness of $M(\xi)$ ensures that the second stage problem is always solvable. However, this situation still fails to satisfy the requirement in [17].

2.3 Quantitative stability

In this subsection, we consider the quantitative stability analysis of problem (1). Denote by $S(P)$ and $\upsilon(P)$ the optimal solution set and optimal value of problem (3), respectively. Similarly, we use $S(Q)$ and $\upsilon(Q)$ to denote the optimal solution set and optimal value of the perturbed problem (3) with probability measure Q. In subsection 2.2, we provide conditions that ensure $S(P)$ and $S(Q)$ are nonempty.

Note the fact that

$$\left| \|a\|^2 - \|b\|^2 \right| = \left|(a-b)^T(a+b)\right| \leq \|a-b\| \left(\|a\| + \|b\|\right)$$

for any $a, b \in \mathbb{R}^n$. By this fact and (2), we have the following estimation:

$$\begin{aligned}
&|f_P(x) - f_Q(x)| \\
&\leq \|\mathbb{E}_P[B(\xi)y^*(x,\xi)] - \mathbb{E}_Q[B(\xi)y^*(x,\xi)]\| \\
&\quad \cdot \big(2\,\|x\| + \|\mathrm{mid}\{l,u,x-(Ax+\mathbb{E}_P[B(\xi)y^*(x,\xi)]+q_1)\}\| \\
&\quad + \|\mathrm{mid}\{l,u,x-(Ax+\mathbb{E}_Q[B(\xi)y^*(x,\xi)]+q_1)\}\|\,\big).
\end{aligned} \tag{11}$$

Firstly, we assume that the support set $\varXi$ is a compact subset in $\mathbb{R}^s$. Then we have

$$\|\mathbb{E}_P[B(\xi)y^*(x,\xi)] - \mathbb{E}_Q[B(\xi)y^*(x,\xi)]\| = \left\| \int_{\varXi} B(\xi)y^*(x,\xi)(P-Q)(d\xi) \right\|.$$

It is known from (10) that

$$|B_i(\xi)y^*(x,\xi)| \leq \|B(\xi)y^*(x,\xi)\| \leq \frac{(R+1)C^2}{\kappa_M(\xi)} \max\{1, \|\xi\|\}^{2\alpha},\ i = 1,2,\cdots,n,$$

where $R = \max_{x\in[l,u]}\{1, \|x\|\}$. Moreover, we have

$$0 < \max_{\xi\in\varXi} \left\{ \frac{(R+1)C^2}{\kappa_M(\xi)} \max\{1, \|\xi\|\}^{2\alpha} \right\} < +\infty$$

because of the compactness of $\varXi$ and the positivity and continuity of $\kappa_M(\xi)$. Therefore, the following inequality holds

$$\left| \int_{\varXi} B_i(\xi)y^*(x,\xi)(P-Q)(d\xi) \right| \leq \max_{\xi\in\varXi} \left\{ \frac{(R+1)C^2}{\kappa_M(\xi)} \max\{1, \|\xi\|\}^{2\alpha} \right\} \mathbb{D}_{TV}(P,Q),$$

for $i = 1,2,\cdots,n$. Thus we obtain

$$\left\| \int_{\varXi} B(\xi)y^*(x,\xi)(P-Q)(d\xi) \right\| \leq \sqrt{n}(R+1)C^2 \max_{\xi\in\varXi} \left\{ \frac{\max\{1, \|\xi\|\}^{2\alpha}}{\kappa_M(\xi)} \right\} \mathbb{D}_{TV}(P,Q).$$

For the second term of the right-hand side of (11), we can bound it above by

$$2\left(4R + \|A\|\,R + (R+1)C^2 \max_{\xi\in\varXi} \left(\frac{\max\{1, \|\xi\|\}^{2\alpha}}{\kappa_M(\xi)} \right) + \|q_1\| \right) := \eta.$$

To sum up, we have the following quantitative estimation.

Lemma 2 *Let Assumptions 1 and 3 hold and $\varXi$ be a compact set. Then there exists a positive constant L_1, such that*

$$\sup_{x\in[l,u]} |f_P(x) - f_Q(x)| \leq L_1 \mathbb{D}_{TV}(P,Q),$$

where $L_1 := \eta\sqrt{n}(R+1)C^2 \max_{\xi\in\varXi} \left(\frac{\{1,\|\xi\|\}^{2\alpha}}{\kappa_M(\xi)} \right)$.

Before establishing the relationship between $S(Q)$ and $S(P)$, we introduce the growth function and its inverse. We call $\psi_P : \mathbb{R}_+ \to \mathbb{R}$ the growth function of problem (3) if

$$\psi_P(\tau) := \min\{f_P(x) : d(x, S(P)) \geq \tau, x \in [l, u]\}.$$

It is not difficult to verify from its definition that $\psi_P(\cdot)$ is nondecreasing and lower semicontinuous. Its inverse function is defined by

$$\psi_P^{-1}(t) = \sup\{\tau \in \mathbb{R}_+ : \psi_P(\tau) \leq t\}, \tag{12}$$

which, of course, is nondecreasing too. For more information, we refer to [19, Example 7.63] and [21].

On the basis of Lemma 2, we immediately obtain the following quantitative description of optimal solution sets.

Theorem 1 *Let Assumptions 1 and 3 hold and the support set Ξ be compact. Then*

$$S(Q) \subseteq S(P) + \psi_P^{-1}(L_1 \mathbb{D}_{TV}(P, Q))\mathbb{B},$$

where L_1 is defined in Lemma 2 and $\mathbb{B}$ is the closed unit ball centered at 0.

Proof A similar proof can be found in [21, Theorem 9]. To keep the paper self-contained, we provide a brief proof. If $S(Q) = \emptyset$, the assertion obviously holds. In the following, we assume $S(Q) \neq \emptyset$. For any $\tilde{x} \in S(Q)$, we have $\upsilon(Q) = f_Q(\tilde{x}) = 0$ and $\upsilon(P) = 0$. Then we have

$$\begin{aligned} L_1 \mathbb{D}_{TV}(P, Q) &= L_1 \mathbb{D}_{TV}(P, Q) + f_Q(\tilde{x}) - \upsilon(P) \\ &\geq f_P(\tilde{x}) - f_Q(\tilde{x}) + f_Q(\tilde{x}) - \upsilon(P) \\ &= f_P(\tilde{x}) - \upsilon(P) \\ &\geq \psi_P(d(\tilde{x}, S(P))). \end{aligned}$$

Thus, we have

$$d(\tilde{x}, S(P)) \leq \psi_P^{-1}(L_1 \mathbb{D}_{TV}(P, Q)).$$

Since $\tilde{x} \in S(Q)$ is selected arbitrarily, we have actually shown that

$$S(Q) \subseteq S(P) + \psi_P^{-1}(L_1 \mathbb{D}_{TV}(P, Q))\mathbb{B}.$$

□

In what follows, we derive the corresponding conclusions without compactness of the support set Ξ by utilizing the conclusion in Lemma 1.

Theorem 2 *Suppose that Assumptions 1 and 3 hold, $P, Q \in \mathcal{P}_{2\alpha+1}(\Xi)$ and $\kappa_M(\xi) \geq \kappa > 0$. Then there exists a positive constant L_2, such that*

$$\sup_{x \in [l,u]} |f_P(x) - f_Q(x)| \leq L_2 \mathbb{D}_{TV}(P, Q)^{\frac{1}{2\alpha+1}}, \tag{13}$$

$$S(Q) \subseteq S(P)) + \psi_P^{-1}(L_2 \mathbb{D}_{TV}(P, Q)^{\frac{1}{2\alpha+1}})\mathbb{B}, \tag{14}$$

when $\mathbb{D}_{TV}(P, Q) + \zeta_{2\alpha+1}(P, Q) \leq 1$.

Proof We know from Lemma 1 that there exists a positive constant $L > 0$, such that

$$\|\mathbb{E}_P[B(\xi)y^*(x,\xi)] - \mathbb{E}_Q[B(\xi)y^*(x,\xi)]\| \leq L\mathbb{D}_{TV}(P,Q)^{\frac{1}{2\alpha+1}}, \tag{15}$$

when $\mathbb{D}_{TV}(P,Q) + \zeta_{2\alpha+1}(P,Q) \leq 1$.

For the second term in the right-hand side of (11), we can bound it above by

$$\begin{aligned}
& 8R + 2\|A\|R + 2\|q_1\| + \frac{(R+1)C^2\big(\mathbb{E}_P[\|\xi\|^{2\alpha}] + \mathbb{E}_Q[\|\xi\|^{2\alpha}] + 2\big)}{\kappa} \\
\leq\ & 8R + 2\|A\|R + 2\|q_1\| + \frac{2(R+1)C^2}{\kappa}(\mathbb{E}_P[\|\xi\|^{2\alpha}] + \alpha\zeta_{2\alpha}(P,Q) + 1) \\
\leq\ & 8R + 2\|A\|R + 2\|q_1\| + \frac{2(R+1)C^2}{\kappa}(\mathbb{E}_P[\|\xi\|^{2\alpha}] + \alpha + 1) := C_1,
\end{aligned} \tag{16}$$

where R is defined as that in Lemma 1 and the second inequality comes from (see [21])

$$\left|\mathbb{E}_Q[\|\xi\|^{2\alpha}] - \mathbb{E}_P[\|\xi\|^{2\alpha}]\right| \leq 2\alpha\zeta_{2\alpha}(P,Q).$$

Combining (15) and (16), and letting $L_2 = LC_1$, we obtain (13). We can derive (14) by using a similar proof as that of Theorem 1, and thus omit the proof. □

Theorems 1 and 2 assert that the solution set of the perturbed problem can be somehow bounded by that of the original problem under specific conditions. In order to quantify it, we adopt a general growth function, instead of imposing a specific growth condition, on the objective function of the original problem. Since the general growth function will vanish at 0, see [21] for details, a sufficiently small perturbation will not change the solution set too much. This stability property is important for both theoretical research and practical calculation. Recall that we say the general growth function ψ_P has the kth order growth for some scalar $k \geq 1$ if $\psi_P(\tau) \geq C\tau^k$ for small $\tau \in \mathbb{R}_+$ and positive constant C. If ψ_P has kth order growth, Theorems 1 and 2 would establish the Hölder continuity of $S(\cdot)$ at P with rate $1/k$.

3 Exponential rate of convergence

In this section, we consider the discrete approximation to problem (1). Assume that, according to the probability distribution P, we have independent and identically distributed samples $\xi^1, \xi^2, \cdots, \xi^K$. Then, for each fixed positive integer K, we have the following discrete approximation to problem (1) with the sample size K, i.e.,

$$\begin{cases} 0 \in Ax + \frac{1}{K}\sum_{i=1}^K (B(\xi^i)y(\xi^i)) + q_1 + \mathcal{N}_{[l,u]}(x), \\ 0 \leq y(\xi^i) \perp M(\xi^i)y(\xi^i) + N(\xi^i)x + q_2(\xi^i) \geq 0, \text{ for } i = 1,2,\cdots,K. \end{cases} \tag{17}$$

In the sequel, we investigate the approximation properties between problems (1) and (17) as K tends to infinity. To this end, we define the discrete approximation distribution P_K with the sample size K by

$$P_K(\xi) = \frac{1}{K}\sum_{i=1}^{K}\delta_{\xi^i}(\xi), \text{ for } \xi \in \Xi,$$

where $\delta_{\xi^i}(\cdot)$ are indicator functions, that is, $\delta_{\xi^i}(\xi) = 1$ if $\xi = \xi^i$; otherwise $\delta_{\xi^i}(\xi) = 0$ for $i = 1, 2, \cdots, K$. Under Assumption 1, we can equivalently rewrite (17) as a minimization problem as follows:

$$\min_{x\in[l,u]} f_{P_K}(x), \tag{18}$$

where f_{P_K} is defined in (2) by substituting P with P_K.

Different from the usual convergence analysis about stochastic variational inequality problems (see for instance [2,3,22]) which does not adopt the residual function, we consider the convergence and exponential rate of convergence between problems (3) and (18).

To investigate the convergence of the optimal solution set of problem (18) to that of problem (3), we need to consider the convergence between $f_{P_K}(x)$ and $f_P(x)$. For this purpose, we first derive the uniform convergence of term $\|\mathbb{E}_P[B(\xi)y^*(x,\xi)] - \mathbb{E}_{P_K}[B(\xi)y^*(x,\xi)]\|$ on $[l,u]$. Thus, we have the following proposition.

Proposition 7 *Suppose that Assumptions 1 and 3 hold, and $P \in \mathcal{P}(\Xi)$ satisfies*

$$\mathbb{E}_P\left[\frac{\|\xi\|^{2\alpha}}{\kappa_M(\xi)}\right] < +\infty. \tag{19}$$

Then

$$\sup_{x\in[l,u]} \|\mathbb{E}_P[B(\xi)y^*(x,\xi)] - \mathbb{E}_{P_K}[B(\xi)y^*(x,\xi)]\| \to 0$$

as $K \to \infty$, with probability one.

Proof It is easy to see from Proposition 1 that $B(\xi)y^*(\cdot,\xi)$ is continuous in $[l,u]$. Moreover, we know from (9) and Assumption 3 that

$$\begin{aligned}\|B(\xi)y^*(x,\xi)\| &\le \frac{1}{\kappa_M(\xi)}\|N(\xi)x + q_2(\xi)\|\,\|B(\xi)\| \\ &\le \frac{(R+1)C^2}{\kappa_M(\xi)}\max\{1,\|\xi\|\}^{2\alpha} \\ &\le \frac{(R+1)C^2}{\kappa_M(\xi)}\left(1+\|\xi\|^{2\alpha}\right). \end{aligned}\tag{20}$$

By (19), we have that the right-hand side of (20) is integrable under probability distribution P. All these arguments ensure the uniform convergence by [23, Theorem 7.53]. ◻

Based on Proposition 7, we immediately obtain the following corollary.

Corollary 1 *Under the same assumptions of Proposition 7, we have that*

$$\lim_{K\to\infty} \sup_{x\in[l,u]} \|\mathbb{E}_{P_K}[B(\xi)y^*(x,\xi)]\| \le \sup_{x\in[l,u]} \|\mathbb{E}_P[B(\xi)y^*(x,\xi)]\| + 1$$

with probability one.

Lemma 3 *Let ψ_P^{-1} be defined in* (12)*. Then for any $\epsilon > 0$, there exists a sufficiently small scalar $\delta > 0$ such that $\psi_P^{-1}(\delta) \le \epsilon$, namely, $\psi_P^{-1}(\delta) \to 0$ as $\delta \to 0$.*

Proof Recall that

$$\psi_P^{-1}(t) = \sup\{\tau \in \mathbb{R}_+ : \psi_P(\tau) \le t\}.$$

For any $\epsilon > 0$, there exists a sufficiently small $\delta > 0$ with $\delta \le \psi_P(\epsilon)$, which implies $\epsilon \ge \psi_P^{-1}(\delta)$. □

Corollary 2 *Let Assumption 1 hold. Then $S(P_K) \neq \emptyset$ for any positive integer K .*

Proof Since P_K is the empirical distribution with finite support set $\{\xi^1, \cdots, \xi^K\}$, we have that

$$\begin{aligned}
&\int_{\Xi} \|B(\xi)\| \max_{J\in\mathcal{J}} \left\|M_{J\times J}^{-1}(\xi)\right\| \|N(\xi)\| \, P_K(d\xi) \\
&= \frac{1}{K}\sum_{i=1}^K \left\|B(\xi^i)\right\| \max_{J\in\mathcal{J}} \left\|M_{J\times J}^{-1}(\xi^i)\right\| \left\|N(\xi^i)\right\| \\
&\le \frac{1}{K}\sum_{i=1}^K \frac{\left\|B(\xi^i)\right\| \left\|N(\xi^i)\right\|}{\kappa_M(\xi^i)} \\
&< +\infty
\end{aligned}$$

for any positive integer K, where the last inequality comes from Assumption 1. Then according to Proposition 3 with $Q = P_K$, $S(P_K)$ is nonempty. □

Theorem 3 *Under assumptions of Proposition 7, we have*

$$d(S(P_K), S(P)) \to 0$$

as $K \to \infty$, with probability one.

Proof Note that $S(P_K)$ is nonempty from Corollary 2. For any $\tilde{x} \in S(P_K)$, we have $\upsilon(P_K) = f_{P_K}(\tilde{x}) = 0$ and $\upsilon(P) = 0$. Then we have

$$\begin{aligned}
\sup_{x\in[l,u]} |f_P(x) - f_{P_K}(x)| &= \sup_{x\in[l,u]} |f_P(x) - f_{P_K}(x)| + f_{P_K}(\tilde{x}) - \upsilon(P) \\
&\ge f_P(\tilde{x}) - f_{P_K}(\tilde{x}) + f_{P_K}(\tilde{x}) - \upsilon(P) \\
&= f_P(\tilde{x}) - \upsilon(P) \\
&\ge \psi_P(d(\tilde{x}, S(P))).
\end{aligned}$$

Thus, we have

$$d(\tilde{x}, S(P)) \leq \psi_P^{-1}\left(\sup_{x\in[l,u]} |f_P(x) - f_{P_K}(x)|\right)$$

for any $\tilde{x} \in S(P_K)$, which implies that

$$d(S(P_K), S(P)) \leq \psi_P^{-1}\left(\sup_{x\in[l,u]} |f_P(x) - f_{P_K}(x)|\right). \tag{21}$$

Therefore, to establish the assertion, we only need to prove

$$\sup_{x\in[l,u]} |f_P(x) - f_{P_K}(x)| \to 0$$

with probability one as $K \to \infty$. We have from (11) that

$$\begin{aligned}
&\sup_{x\in[l,u]} |f_P(x) - f_{P_K}(x)| \leq \\
&\qquad \sup_{x\in[l,u]} \|\mathbb{E}_P[B(\xi)y^*(x,\xi)] - \mathbb{E}_{P_K}[B(\xi)y^*(x,\xi)]\| \cdot \theta(P_K, P),
\end{aligned}$$

where

$$\begin{aligned}
\theta(P_K, P) =& 8R + 2R\|A\| + 2\|q_1\| \\
&+ \sup_{x\in[l,u]} \|\mathbb{E}_{P_K}[B(\xi)y^*(x,\xi)]\| + \sup_{x\in[l,u]} \|\mathbb{E}_P[B(\xi)y^*(x,\xi)]\|.
\end{aligned} \tag{22}$$

Then, we obtain

$$\begin{aligned}
&\lim_{K\to\infty} \sup_{x\in[l,u]} |f_P(x) - f_{P_K}(x)| \leq \\
&\quad \lim_{K\to\infty} \sup_{x\in[l,u]} \|\mathbb{E}_P[B(\xi)y^*(x,\xi)] - \mathbb{E}_{P_K}[B(\xi)y^*(x,\xi)]\| \cdot \lim_{K\to\infty} \theta(P_K, P).
\end{aligned}$$

It can be deduced from Proposition 7 and Corollary 1 that

$$\lim_{K\to\infty} \sup_{x\in[l,u]} \|\mathbb{E}_P[B(\xi)y^*(x,\xi)] - \mathbb{E}_{P_K}[B(\xi)y^*(x,\xi)]\| = 0$$

and

$$\lim_{K\to\infty} \sup_{x\in[l,u]} \|\mathbb{E}_{P_K}[B(\xi)y^*(x,\xi)]\| \leq \sup_{x\in[l,u]} \|\mathbb{E}_P[B(\xi)y^*(x,\xi)]\| + 1$$

with probability one, respectively. The second assertion above indicates that

$$\lim_{K\to\infty} \theta(P_K, P) \leq \lambda(P)$$

with probability one, where

$$\lambda(P) = 8R + 2R\|A\| + 2\|q_1\| + 2\sup_{x\in[l,u]} \|\mathbb{E}_P[B(\xi)y^*(x,\xi)]\| + 1. \tag{23}$$

All these imply that

$$\lim_{K\to\infty} \sup_{x\in[l,u]} |f_P(x) - f_{P_K}(x)| = 0$$

with probability one. Due to Lemma 3, we obtain

$$\lim_{K\to\infty} \psi_P^{-1} \left(\sup_{x\in[l,u]} |f_P(x) - f_{P_K}(x)| \right) = 0$$

with probability one, which completes the proof. □

Now we derive the exponential rate of convergence of the SAA using similar conditions in [24]. The authors of [24] studied the uniformly exponential convergence of the SAA for stochastic mathematical programs with variational constraints through the Cramér's Large Deviation Theorem. To derive the exponential rate of convergence of the SAA for the two-stage stochastic linear variational inequality problem, we need that $[B(\xi)y^*(x,\xi)]$ is Lipschitz continuous w.r.t. x, that is,

$$\begin{aligned} \|B(\xi)y^*(x_1,\xi) - B(\xi)y^*(x_2,\xi)\| &\le \|B(\xi)\| \, \|N(\xi)\| \, \|x_1 - x_2\| / \kappa_M(\xi) \\ &= C(\xi) \, \|x_1 - x_2\| \end{aligned}$$

where $C(\xi) = \|B(\xi)\| \, \|N(\xi)\| / \kappa_M(\xi)$, which is ensured by Assumption 1.

To establish the exponential rate of convergence, similar to that in [24], we need the following assumptions.

Assumption 4 *Let the following assertions hold:*

(i) For each $x \in [l,u]$, the moment generating functions of random variables $[B(\xi)y^*(x,\xi)]_i - (\mathbb{E}_P[B(\xi)y^*(x,\xi)])_i$, i.e.,

$$\mathbb{E}_P[\exp(t([B(\xi)y^*(x,\xi)]_i - (\mathbb{E}_P[B(\xi)y^*(x,\xi)])_i))]$$

for $i = 1, 2, \cdots, n$, are finite valued for each t in a neighborhood of zero;

(ii) The moment generating function of $C(\xi)$, i.e.,

$$\mathbb{E}_P[\exp(tC(\xi))]$$

is finite valued for each t in a neighborhood of zero.

Proposition 8 *Let Assumptions 1 and 4 hold. Then for any $\epsilon > 0$, there exist two positive scalars $L(\epsilon)$ and $\beta(\epsilon)$ which depend only on ϵ, such that*

$$\mathbb{P}\left\{ \sup_{x\in[l,u]} \|\mathbb{E}_{P_K}[B(\xi)y^*(x,\xi)] - \mathbb{E}_P[B(\xi)y^*(x,\xi)]\| \ge \epsilon \right\} \le L(\epsilon)\exp(-K\beta(\epsilon)).$$

This proposition can be directly obtained from [24, Theorem 5.1]. We thus omit the proof here.

From Proposition 8, we can immediately obtain the following exponential rate of convergence about the optimal solution set.

Theorem 4 *Let Assumptions 1 and 4 hold. Then, for any $\epsilon > 0$, there exist two positive scalars $\bar{L}(\epsilon)$ and $\bar{\beta}(\epsilon)$, such that*

$$\mathbb{P}\left\{d(S(P_K), S(P)) \geq \epsilon\right\} \leq \bar{L}(\epsilon)\exp(-K\bar{\beta}(\epsilon)).$$

Proof We have from (21) the following estimation:

$$\begin{aligned}\mathbb{P}\{d(S(P_K), S(P)) \geq \epsilon\} &\leq \mathbb{P}\left\{\psi_P^{-1}\left(\sup_{x\in[l,u]}|f_P(x) - f_{P_K}(x)|\right) \geq \epsilon\right\}\\ &\leq \mathbb{P}\left\{\sup_{x\in[l,u]}|f_P(x) - f_{P_K}(x)| \geq \psi_P(\epsilon)\right\}.\end{aligned}$$

The second inequality follows from the nondecreasing property of ψ_P.

We know from Proposition 8 that

$$\mathbb{P}\left\{\sup_{x\in[l,u]}\|\mathbb{E}_{P_K}[B(\xi)y^*(x,\xi)] - \mathbb{E}_P[B(\xi)y^*(x,\xi)]\| < 1\right\} \geq 1-L(1)\exp(-K\beta(1)),$$

which implies

$$\begin{aligned}&\mathbb{P}\left\{\sup_{x\in[l,u]}\|\mathbb{E}_{P_K}[B(\xi)y^*(x,\xi)]\| < \sup_{x\in[l,u]}\|\mathbb{E}_P[B(\xi)y^*(x,\xi)]\| + 1\right\}\\ &\geq 1 - L(1)\exp(-K\beta(1)).\end{aligned}$$

In addition, we have that

$$\begin{aligned}&\sup_{x\in[l,u]}|f_P(x) - f_{P_K}(x)| \leq\\ &\quad\sup_{x\in[l,u]}\|\mathbb{E}_P[B(\xi)y^*(x,\xi)] - \mathbb{E}_{P_K}[B(\xi)y^*(x,\xi)]\| \cdot \theta(P_K, P),\end{aligned}$$

where $\theta(P_K, P)$ is defined in (22). Therefore, we obtain

$$\mathbb{P}\left\{\theta(P_K, P) < \lambda(P)\right\} \geq 1 - L(1)\exp(-K\beta(1)),$$

where $\lambda(P)$ is defined in (23). Thus, we continue

$$\begin{aligned}&\mathbb{P}\left\{\sup_{x\in[l,u]}|f_P(x) - f_{P_K}(x)| \geq \psi_P(\epsilon)\right\}\\ &\leq \mathbb{P}\left\{\sup_{x\in[l,u]}\|\mathbb{E}_P[B(\xi)y^*(x,\xi)] - \mathbb{E}_{P_K}[B(\xi)y^*(x,\xi)]\| \cdot \theta(P_K, P) \geq \psi_P(\epsilon)\right\}\\ &\leq L(1)\exp(-K\beta(1)) +\\ &\quad\mathbb{P}\left\{\sup_{x\in[l,u]}\|\mathbb{E}_P[B(\xi)y^*(x,\xi)] - \mathbb{E}_{P_K}[B(\xi)y^*(x,\xi)]\| \cdot \lambda(P) \geq \psi_P(\epsilon)\right\}\\ &\leq L(1)\exp(-K\beta(1)) + L\left(\frac{\psi_P(\epsilon)}{\lambda(P)}\right)\exp\left(-K\beta\left(\frac{\psi_P(\epsilon)}{\lambda(P)}\right)\right),\end{aligned}$$

where the third inequality comes from Proposition 8.

Letting

$$\bar{L}(\epsilon) := L(1) + L\left(\frac{\psi_P(\epsilon)}{\lambda(P)}\right)$$

and

$$\bar{\beta}(\epsilon) := \min\left\{\beta(1), \beta\left(\frac{\psi_P(\epsilon)}{\lambda(P)}\right)\right\},$$

we complete the proof. □

In this section, we study the discrete approximation properties of problem (1) under mild conditions. The convergence of the SAA is derived in Theorem 3. However, this result did not address an important issue which is interesting from both the theoretical and computational points of view. That is, what is the rate of convergence or how large the sample size should be to achieve a desired accuracy of SAA estimators. We supplement it in Theorem 4 under ordinary assumptions. These estimates provide an important insight into the theoretical complexity and practical application of the considered problem (1).

4 Numerical results

To illustrate the application of the two-stage stochastic linear variational inequality problem (1) and to verify the obtained convergence results, we consider in this section a multi-player non-cooperative two-stage game problem (see also [2,3] for the two-players case) and its numerical solution. There is a significant amount of recent research on this topic. For example, [11] investigated the two-stage game wherein each player is risk-averse and solved a rival-parameterized stochastic program with quadratic recourse. The convergence results for different versions of the best-response schemes are discussed. [9] considered a stochastic Nash game where each player minimizes a parameterized expectation-valued convex objective function by proposing three inexact proximal best-response schemes. Different from those in [9,11] where the Nash equilibrium point is determined by (inexact) best-response schemes, we employ the PHM to solve the discrete two-stage stochastic variational inequality problem (17).

4.1 A multi-player non-cooperative two-stage game

Two-stage stochastic variational inequality problems have many practical applications (see [1]). Here we consider the multi-player (say $\mathcal{I}$ players) non-cooperative two-stage game. It can be described in the form of the two-stage stochastic variational inequality problem (1). Let $(x_1, y_1(\cdot))$, $(x_2, y_2(\cdot)), \cdots,$ $(x_{\mathcal{I}}, y_{\mathcal{I}}(\cdot)) \in \mathbb{R}^n \times \mathcal{Y}$ denote the decisions of player 1 to player $\mathcal{I}$ in the two-stage stochastic game, respectively. We use x_{-i} to denote $\{x_j\}_{j\neq i}$ and so does y_{-i}. $\theta_i : \mathbb{R}^{n\mathcal{I}} \to \mathbb{R}$ is the cost function of player i in the first stage and

$\phi_i : \mathbb{R}^{n\mathcal{I}} \times \mathcal{Y}^{\mathcal{I}} \times \Xi \to \mathbb{R}$ is the cost function of player i in the second stage. Then, to minimize his total cost, the player i $(i = 1, 2, \cdots, \mathcal{I})$ will make a decision through solving the following two-stage stochastic optimization problem:

$$\min_{x_i \in [l_i, u_i]} \theta_i(x_i, x_{-i}) + \mathbb{E}_P[\varphi_i(x_i, x_{-i}, y_{-i}(\xi), \xi)], \tag{24}$$

for $l_i < u_i$ and $l_i, u_i \in \mathbb{R}^n$, $i = 1, 2, \cdots, \mathcal{I}$, where $\varphi_i(x_i, x_{-i}, y_{-i}(\xi), \xi)$ is defined by

$$\varphi_i(x_i, x_{-i}, y_{-i}(\xi), \xi) = \min_{y_i \geq 0} \phi_i(x_i, x_{-i}, y_i, y_{-i}(\xi), \xi). \tag{25}$$

We know that a two-stage stochastic programming problem can be equivalently reformulated as a two-stage variational inequality problem from the first order optimality necessary conditions. Therefore, we consider the optimality condition of the two-stage stochastic program (24)-(25). To simplify the formulation, we assume that $\theta_i(\cdot, x_{-i})$ is differentiable w.r.t. x_i and $\phi_i(x_i, x_{-i}, \cdot, y_{-i}, \xi)$ is differentiable w.r.t. y_i. In addition, $\varphi_i(\cdot, x_{-i}, y_{-i}(\xi), \xi)$ is differentiable and Lipschitz continuous with some integrable Lipschitz constant w.r.t. x_i. Then, we know from [23, Theorem 7.49] that

$$\nabla_{x_i} \mathbb{E}_P[\varphi_i(x_i, x_{-i}, y_{-i}(\xi), \xi)] = \mathbb{E}_P[\nabla_{x_i} \varphi_i(x_i, x_{-i}, y_{-i}(\xi), \xi)].$$

Finally, we obtain the equivalent form of problem (24)-(25) as

$$\begin{cases} 0 \in \nabla_{x_i} \theta_i(x_i, x_{-i}) + \mathbb{E}_P[\nabla_{x_i} \varphi_i(x_i, x_{-i}, y_{-i}(\xi), \xi)] + \mathcal{N}_{[l_i, u_i]}(x_i) \\ 0 \leq y_i \perp \nabla_{y_i} \phi_i(x_i, x_{-i}, y_i, y_{-i}, \xi) \geq 0, \text{ for a.e. } \xi \in \Xi, \end{cases} \tag{26}$$

for $i = 1, 2, \cdots, \mathcal{I}$.

To satisfy the above conditions and to obtain concrete numerical results, we consider a two-stage stochastic quadratic programming problem. Specifically, we define

$$\theta_i(x_i, x_{-i}) = \frac{1}{2} x_i^T H_i x_i + b_i^T x_i + \sum_{j \neq i} x_i^T P_j x_j$$

and

$$\phi_i(x_i, x_{-i}, y_i, y_{-i}, \xi) = \frac{1}{2} y_i^T Q_i(\xi) y_i + c_i(\xi)^T y_i + \sum_{j=1}^{I} y_i^T S_{ij}(\xi) x_j + \sum_{j \neq i} y_i^T O_j(\xi) y_j(\xi),$$

where $H_i, P_i \in \mathbb{R}^{n \times n}$, $S_{ij} : \Xi \to \mathbb{R}^{m \times n}$, $O_i : \Xi \to \mathbb{R}^{m \times m}$, $Q_i : \Xi \to \mathbb{R}^{m \times m}$, $b_i \in \mathbb{R}^n$, $c_i : \Xi \to \mathbb{R}^m$ for $i, j = 1, 2, \cdots, \mathcal{I}$.

With the above notation, we can rewrite problem (26) as the following large-scale two-stage stochastic linear variational inequality problem (see [2]):

$$\begin{cases} 0 \in Ax + \mathbb{E}_P[B(\xi) y(\xi)] + q_1 + \mathcal{N}_{[l,u]}(x), \\ 0 \leq y(\xi) \perp M(\xi) y(\xi) + N(\xi) x + q_2(\xi) \geq 0, \text{ for a.e. } \xi \in \Xi, \end{cases} \tag{27}$$

where

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_{\mathcal{I}} \end{pmatrix}, \quad y(\xi) = \begin{pmatrix} y_1(\xi) \\ \vdots \\ y_{\mathcal{I}}(\xi) \end{pmatrix}, q_1 = \begin{pmatrix} b_1 \\ \vdots \\ b_{\mathcal{I}} \end{pmatrix},$$

$$q_2(\xi) = \begin{pmatrix} c_1(\xi) \\ \vdots \\ c_{\mathcal{I}}(\xi) \end{pmatrix}, \quad l = \begin{pmatrix} l_1 \\ \vdots \\ l_{\mathcal{I}} \end{pmatrix}, \quad u = \begin{pmatrix} u_1 \\ \vdots \\ u_{\mathcal{I}} \end{pmatrix},$$

$$A = \begin{pmatrix} H_1 & P_2 & \cdots & P_{\mathcal{I}} \\ P_1 & H_2 & \cdots & P_{\mathcal{I}} \\ \vdots & \vdots & \ddots & \vdots \\ P_1 & P_2 & \cdots & H_{\mathcal{I}} \end{pmatrix}, \quad B(\xi) = \begin{pmatrix} S_{11}^T(\xi) & 0 & \cdots & 0 \\ 0 & S_{22}^T(\xi) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & S_{\mathcal{I}\mathcal{I}}^T(\xi) \end{pmatrix},$$

$$M(\xi) = \begin{pmatrix} Q_1(\xi) & O_2(\xi) & \cdots & O_{\mathcal{I}}(\xi) \\ O_1(\xi) & Q_2(\xi) & \cdots & O_{\mathcal{I}}(\xi) \\ \vdots & \vdots & \ddots & \vdots \\ O_1(\xi) & O_2(\xi) & \cdots & Q_{\mathcal{I}}(\xi) \end{pmatrix}, N(\xi) = \begin{pmatrix} S_{11}(\xi) & S_{12}(\xi) & \cdots & S_{1\mathcal{I}}(\xi) \\ S_{21}(\xi) & S_{22}(\xi) & \cdots & S_{2\mathcal{I}}(\xi) \\ \vdots & \vdots & \ddots & \vdots \\ S_{\mathcal{I}1}(\xi) & S_{\mathcal{I}2}(\xi) & \cdots & S_{\mathcal{I}\mathcal{I}}(\xi) \end{pmatrix}.$$

A well-known algorithm for solving two-stage stochastic variational inequality problems is PHM, see [17,18,20], which is convergent for monotone problems. The main idea of this algorithm is to construct a nonanticipative first stage solution through solving several discrete problems corresponding to individual scenarios. Let $\xi^1, \xi^2, \cdots, \xi^K$ be K samples or scenarios, and PHM can be stated as follows.

Algorithm 1 (PHM to solve problem (27)**)**

Step 0: *Choose initial points:* $\bar{x}_0$, $x_0^k = \bar{x}_0$, y_0^k, w_0^k *with* $\sum_{k=1}^K w_0^k = 0$, *and for* $k = 1, 2, \cdots, K$. *Let* $r > 0$ *and set* $i = 0$.
Step 1: *For* $k = 1, 2, \cdots, K$, *solve the following two-stage mixed problem:*

$$\begin{cases} 0 \in Ax + B(\xi^k)y + q_1 + w_i^k + r(x - x_i^k) + \mathcal{N}_{[l,u]}(x), \\ 0 \le y \perp M(\xi^k)y + N(\xi^k)x + q_2(\xi^k) + r(y - y_i^k) \ge 0. \end{cases} \tag{28}$$

The obtained solution is denoted by $(\hat{x}_i^k, \hat{y}_i^k)$;
Step 2: *Let* $\bar{x}_{i+1} = \frac{1}{K}\sum_{k=1}^K \hat{x}_i^k$. *Then, for* $k = 1, 2, \cdots, K$, *set* $x_{i+1}^k = \bar{x}_{i+1}$, $y_{i+1}^k = \hat{y}_i^k$ *and* $w_{i+1}^k = w_i^k + r(\hat{x}_i^k - \bar{x}_{i+1})$. *If a termination criterion is satisfied, stop. Otherwise, let* $i = i + 1$ *and go back to* ***Step 1****.*

A termination criterion can be chosen as

$$\frac{1}{K}\sum_{k=1}^K \left\| x_i^k - \mathrm{mid}\left\{ l, u, x_i^k - \left(Ax_i^k + \frac{1}{K}\sum_{k=1}^K B(\xi^k)y_i^k + q_1 \right) \right\} \right\| \le \epsilon, \tag{29}$$

where ϵ is a sufficiently small positive number.

4.2 Parameter settings and numerical results

We consider a 3-player two-stage non-cooperative game with $n = 4,\ m = 4$. We adopt the following stopping criterion for PHM: Either the residual in (29) is less than or equal to 10^{-5} or the iteration number i attains 6000. Arbitrarily generate $\hat{H}_i \in \mathbb{R}^{n\times n}$, $\hat{Q}_i \in \mathbb{R}^{m\times m}$, $\hat{P}_i \in \mathbb{R}^{n\times n}$, $\hat{S}_{ij} \in \mathbb{R}^{m\times n}$, $\hat{O}_i \in \mathbb{R}^{m\times m}$ with entries choosing from $[-1, 1]$ and $b_i \in \mathbb{R}^n$, $c_i \in \mathbb{R}^m$, for $i, j = 1, 2, 3$. Let $\xi = (\xi_1, \xi_2, \cdots, \xi_{18})$ be the random vector which follows a uniform distribution on the support set $[0, 1]^{18}$. Then we set $\hat{S}_{11}(\xi) = \xi_1\hat{S}_{11}$, $\hat{S}_{12}(\xi) = \xi_2\hat{S}_{12}$, $\hat{S}_{13}(\xi) = \xi_3\hat{S}_{13}$, $\hat{S}_{21}(\xi) = \xi_4\hat{S}_{21}$, $\hat{S}_{22}(\xi) = \xi_5\hat{S}_{22}$, $\hat{S}_{23}(\xi) = \xi_6\hat{S}_{23}$, $\hat{S}_{31}(\xi) = \xi_7\hat{S}_{31}$, $\hat{S}_{32}(\xi) = \xi_8\hat{S}_{32}$, $\hat{S}_{33}(\xi) = \xi_9\hat{S}_{33}$, $\hat{O}_1(\xi) = \xi_{10}\hat{O}_1$, $\hat{O}_2(\xi) = \xi_{11}\hat{O}_2$, $\hat{O}_3(\xi) = \xi_{12}\hat{O}_3$, $\hat{Q}_1(\xi) = \xi_{13}\hat{Q}_1$, $\hat{Q}_2(\xi) = \xi_{14}\hat{Q}_2$, $\hat{Q}_3(\xi) = \xi_{15}\hat{Q}_3$, $c_1(\xi) = \xi_{16}c_1$, $c_2(\xi) = \xi_{17}c_2$ and $c_3(\xi) = \xi_{18}c_3$. The main reason to choose the above random parameters is to satisfy Assumption 3, which is needed in Theorems 1 and 3. Meanwhile, there are plenty of existing works and applications where the parameters are assumed to be affinely linear w.r.t. ξ, see for example [21].

To ensure the positive definiteness of coefficient matrices in problem (27), we construct those matrices as follows:

$$A = \begin{pmatrix} \hat{H}_1 & \hat{P}_2 & \hat{P}_3 \\ \hat{P}_1 & \hat{H}_2 & \hat{P}_3 \\ \hat{P}_1 & \hat{P}_2 & \hat{H}_3 \end{pmatrix} + \gamma I_{3n}, \qquad B(\xi) = \begin{pmatrix} \hat{S}_{11}^T(\xi) & 0 & 0 \\ 0 & \hat{S}_{22}^T(\xi) & 0 \\ 0 & 0 & \hat{S}_{33}^T(\xi) \end{pmatrix},$$

$$M(\xi) = \begin{pmatrix} \hat{Q}_1(\xi) & \hat{O}_2(\xi) & \hat{O}_3(\xi) \\ \hat{O}_1(\xi) & \hat{Q}_2(\xi) & \hat{O}_3(\xi) \\ \hat{O}_1(\xi) & \hat{O}_2(\xi) & \hat{Q}_3(\xi) \end{pmatrix} + \gamma I_{3m}, \ N(\xi) = \begin{pmatrix} \hat{S}_{11}(\xi) & \hat{S}_{12}(\xi) & \hat{S}_{13}(\xi) \\ \hat{S}_{21}(\xi) & \hat{S}_{22}(\xi) & \hat{S}_{23}(\xi) \\ \hat{S}_{31}(\xi) & \hat{S}_{32}(\xi) & \hat{S}_{33}(\xi) \end{pmatrix},$$

$$q_1 = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}, \qquad q_2(\xi) = \begin{pmatrix} c_1(\xi) \\ c_2(\xi) \\ c_3(\xi) \end{pmatrix},$$

where $\gamma = 3(m + n)$, I_{3n} and I_{3m} stand for the identity matrices in $\mathbb{R}^{3n\times 3n}$ and $\mathbb{R}^{3m\times 3m}$, respectively. Obviously, the above setting guarantees that Assumption 2 holds for any $\xi \in [0, 1]^{18}$, which is sufficient for the convergence of the PHM. Due to the affine linearity of all the above coefficients, Assumption 3 holds with $\alpha = 1$. Moreover, we adopt the uniform distribution here that must satisfy (19) in Proposition 7. Therefore, Theorems 1 and 3 hold in our specific settings.

From (i) of Proposition 1, the solution of the second stage satisfies

$$\|y^*(x, \xi)\| \le \|W(x, \xi)\| \left(\|N(\xi)\| \, \|x\| + \|q_2(\xi)\|\right) \le \Gamma$$

uniformly for any $\xi \in [0, 1]^{18}$ and some positive number Γ. This implies that we can employ the homotopy-smoothing method for box-constrained variational inequalities (see [5]) to solve the two-stage mixed problem (28) in **Step 1**.

With the above detailed parameter selection and the solution method in **Step 1**, we can then solve the concrete 3-player two-stage non-cooperative game problem. We show in Figure 1 the box plot for each component of the first stage decision variable x w.r.t. the number of samples. Since our parameter setting satisfies Assumption 2, there exist a unique solution for both the original problem and its SAA problem (see [3]). As we discussed before, Theorem 3 holds in our setting. For each sample size $K = 10, 50, 200, 500, 1000, 2000, 4000$, we solve 100 randomly generated problems and draw the empirical distribution of the solutions in Figure 1. The 12 plots in Figure 1 show the convergence of the SAA problem (17) by adopting the hybrid algorithm combining PHM [18] and the homotopy-smoothing method [5] as the sample size goes to infinity. Actually, we know from Theorem 3 and the uniqueness of solution that the SAA solutions will converge to the true solution with probability one.

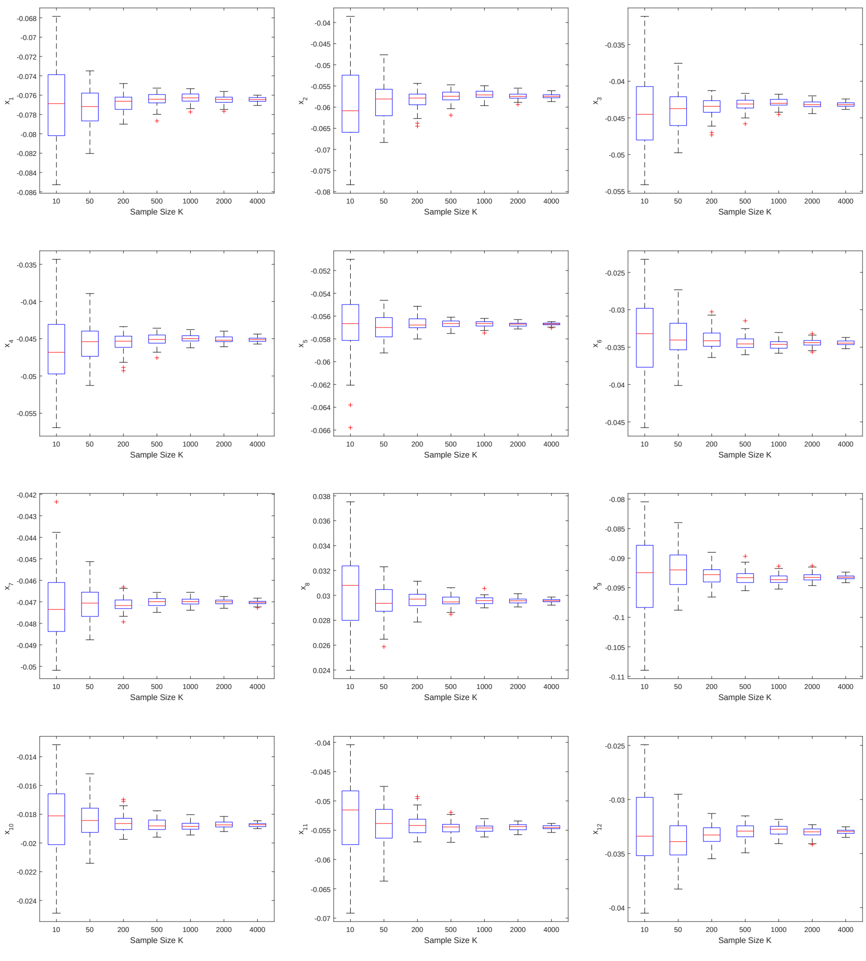

Fig. 1: The box plots for components of x

Now we numerically verify the quantitative stability results in section 2 to this example. For this purpose, we assume that the original probability distribution P is the uniform distribution on interval $[0,1]^{18}$. The perturbed distribution Q_ν ($\nu \in \mathbb{N}$) is the uniform distribution with the support set being $\left[0, \frac{\nu}{\nu+1}\right]^{18}$, that is, the probability for taking values in $[0,1]^{18} \setminus \left[0, \frac{\nu}{\nu+1}\right]^{18}$ is zero. Then, we have

$$\begin{aligned}\mathbb{D}_{TV}(P, Q_\nu) &= \sup_{h \in \mathcal{G}_{TV}} \left\{ \int_{\left[0, \frac{\nu}{\nu+1}\right]^{18}} h(\xi) \left(\left(\frac{\nu+1}{\nu} \right)^{18} - 1 \right) d\xi - \int_{[0,1]^{18} \setminus \left[0, \frac{\nu}{\nu+1}\right]^{18}} h(\xi) d\xi \right\} \\ &= 2 \left[1 - \left(\frac{\nu}{\nu+1} \right)^{18} \right]. \end{aligned} \tag{30}$$

Here the optimal element in $\mathcal{G}_{TV}$ is

$$h(\xi) = \begin{cases} 1, & \xi \in \left[0, \frac{\nu}{\nu+1}\right]^{18}; \\ -1, & \xi \in [0,1]^{18} \setminus \left[0, \frac{\nu}{\nu+1}\right]^{18}. \end{cases}$$

Therefore, $\mathbb{D}_{TV}(P, Q_\nu) \to 0$ as $\nu \to +\infty$. In what follows, we fix the number of scenarios at $K = 5000$ and use the sample approximation problem to approximate the original problem. Let $\nu = 1, 2, 3, 4, 5, 6, 7$, we use PHM to solve the original problem under P and the corresponding problem under perturbed distribution Q_ν, respectively. Since Assumption 2 holds, there always exists a unique solution for the original problem under P, as well as the problem under the perturbation Q_ν.

We calculate the distance between the unique solution x^* under probability distribution P and the unique solution x^*_ν under probability distribution Q_ν. It is known from Theorem 1 that

$$\|x^* - x^*_\nu\| \leq \psi_P^{-1}(L_1 \mathbb{D}_{TV}(P, Q_\nu)) \tag{31}$$

for some positive constant L_1. Note that ψ_P^{-1} is lower semicontinuous and nondecreasing, and vanishes at 0. Specially, under our specific setting, we know from

$$\psi_P(\tau) = \min\{f_P(x) = f_P(x) - f_P(x^*) : d(x, x^*) \geq \tau, x \in [l, u]\},$$

where $f_P(x^*) = 0$, and the continuity of $f_P(x)$ w.r.t. x that ψ_P is continuous at 0. Moreover, $\psi_P(\tau) > 0$ for any $\tau > 0$ due to the uniqueness of solutions. Its inverse function is defined by (12), that is,

$$\psi_P^{-1}(t) = \sup\{\tau \in \mathbb{R}_+ : \psi_P(\tau) \leq t\}$$

is continuous at $t = 0$, see Lemma 3.

Based on the above discussion, we have from (31) that $\|x^* - x^*_\nu\|$ should converge to 0 as $\nu \to \infty$. Table 1 shows this kind of convergence. We can see from Table 1 that the distance between x^* and x^*_ν monotonically decreases with the increase of ν. These results perfectly illustrate and support the quantitative analysis results in section 2.

Table 1: The distance between the pairing solutions under P and Q_ν

ν	1	2	3	4	5	6	7
$\|x^* - x^*_\nu\|$	1.33e-2	1.00e-2	0.77e-2	0.64e-2	0.54e-2	0.49e-2	0.41e-2

5 Concluding remarks

In this paper, we study a class of two-stage stochastic linear variational inequality problems through the residual minimization problem (3). The quantitative stability and convergence analysis are conducted with respect to problem (3). Specifically, we first provide sufficient conditions for the existence of solutions of both the original problem and the perturbed problems. Next we conduct the quantitative stability analysis under the total variation metric, and further investigate the convergence of discrete approximations of the two-stage linear stochastic variational inequality problem. Finally, by a 3-player two-stage non-cooperative game problem, we numerically illustrate our convergence conclusion and quantitative stability results.

There are still a few issues that are worth for further investigation. For example, quantitative stability analysis for multi-stage stochastic linear/nonlinear variational inequality problems.

Acknowledgements The authors thank Dr. Hailin Sun for his help in the numerical experiment. They would like to thank the editor and three anonymous referees for their insightful comments which help us to improve the content and presentation of this manuscript.

References

1. Chen, X., Pong, T.K., Wets, R.J.: Two-stage stochastic variational inequalities: an ERM-solution procedure. Math. Program. **165**, 71–111 (2017)
2. Chen, X., Shapiro, A., Sun, H.: Convergence analysis of sample average approximation of two-state stochastic generalized equations. SIAM J. Optim. **29**, 135–161 (2019)
3. Chen, X., Sun, H., Xu, H.: Discrete approximation of two-stage stochastic and distributionally robust linear complementarity problems. Math. Program. **177**, 255–289 (2019)
4. Chen, X., Xiang, S.: Newton iterations in implicit time-stepping scheme for differential linear complementarity systems. Math. Program. **138**, 579–606 (2013)
5. Chen, X., Ye, Y.: On homotopy-smoothing methods for box-constrained variational inequalities. SIAM J. Control Optim. **37**, 589–616 (1999)
6. Cottle, R.W., Pang, J.S., Stone, R.E.: The Linear Complementarity Problem. SIAM, Philadelphia (2009)
7. Facchinei, F., Pang, J.S.: Finite-dimensional Variational Inequalities and Complementarity Problems. Springer-Verlag, New York (2003)
8. King, A.J., Rockafellar, R.T.: Sensitivity analysis for nonsmooth generalized equations. Math. Program. **55**, 193–212 (1992)

9. Lei, J., Shanbhag, U.V., Pang, J.S., Sen, S.: On synchronous, asynchronous, and randomized best-response schemes for computing equilibria in stochastic Nash games. Math. Oper. Res. 45, 157-190 (2019)
10. Luo, Z.Q., Pang, J.S., Ralph, D.: Mathematical Programs with Equilibrium Constraints. Cambridge University Press, Cambridge (1996)
11. Pang, J.S., Sen, S., Shanbhag, U.V.: Two-stage non-cooperative games with risk-averse players. Math. Program. **165**, 235–290 (2017)
12. Rachev, S.T., Klebanov, L., Stoyanov, S.V., Fabozzi, F.: The methods of distances in the theory of probability and statistics. Springer, New York (2013)
13. Ravat, U., Shanbhag, U.V.: On the characterization of solution sets of smooth and nonsmooth convex stochastic nash games. SIAM J. Optim. **21**, 1168–1199 (2011)
14. Ravat, U., Shanbhag, U.V.: On the existence of solutions to stochastic quasi-variational inequality and complementarity problems. Math. Program. **165**, 291–330 (2017)
15. Robinson, S.M.: Generalized equations and their solutions, part i: Basic theory. In: Point-to-Set Maps and Mathematical Programming, pp. 128–141. Springer (1979)
16. Robinson, S.M.: Strongly regular generalized equations. Math. Oper. Res. **5**, 43–62 (1980)
17. Rockafellar, R.T., Sun, J.: Solving monotone stochastic variational inequalities and complementarity problems by progressive hedging. Math. Program. **174**, 453–471 (2019)
18. Rockafellar, R.T., Wets, R.J.: Scenarios and policy aggregation in optimization under uncertainty. Math. Oper. Res. **16**, 119–147 (1991)
19. Rockafellar, R.T., Wets, R.J.: Variational Analysis. Springer-Verlag, Berlin (2009)
20. Rockafellar, R.T., Wets, R.J.: Stochastic variational inequalities: single-stage to multistage. Math. Program. **165**, 331–360 (2016)
21. Römisch, W.: Stability of stochastic programming problems. Handb. Oper. Res. Manag. Sci. **10**, 483–554 (2003)
22. Shapiro, A.: Monte carlo sampling methods. Handb. Oper. Res. Manag. Sci. **10**, 353–425 (2003)
23. Shapiro, A., Dentcheva, D., Ruszczyński, A.: Lectures on Stochastic Programming: Modeling and Theory. SIAM, Philadelphia (2014)
24. Shapiro, A., Xu, H.: Stochastic mathematical programs with equilibrium constraints, modelling and sample average approximation. Optimization **57**, 395–418 (2008)
25. Zhang, J., Xu, H., Zhang, L.: Quantitative stability analysis for distributionally robust optimization with moment constraints. SIAM J. Optim. **26**, 1855–1882 (2016)